DOCUMENT RESUME

ED 032 119 PS 002 108

By-Osser, Harry; And Others
A Study of the Communicative Abilities of Disadvantaged Children. Final Report.
Johns Hopkins Univ., Baltimore, Md. School of Medicine.
Spons Agency-Office of Economic Opportunity, Washington, D.C.
Report No-OEO-2402
Pub Date 30 Jan 68
Note-45p.
EDRS Price MF-$0.25 HC-$2.35
Descriptors-Caucasians, *Child Language, *Dialects, Language Development, Listening Comprehension, Models, Negroes, Nonstandard Dialects, Preschool Children, Psycholinguistics, Sentence Structure, Structural Analysis, *Structural Linguistics, Syntax

The purpose of this series of four studies was to precisely describe the code and dialect features of the speech of both lower class Negro children and middle class white children. In the first study, 16 white middle class (WMC) children were compared to 16 Negro lower class (NLC) children on both an imitation and a comprehension task. The WMC subjects scored significantly higher on both tasks, even after the scores of the NLC subjects on the imitation tasks were improved by adjusting them for differences of dialect in the children's responses. No adjustment, however, was made for the administering of both tasks in standard English. The second study took free speech samples from 20 NLC and 20 WMC 5-year-olds to discover any possible linguistic code variations between groups. The WMC subjects showed a significantly superior range of syntactic structures, but there was no significant difference between groups in the use of specific types of complex sentences. The same free speech samples were then analyzed in the third study to find and list examples of nonstandard dialect variations of NLC. The fourth study developed a psycholinguistic model for measuring syntactic complexity in bot h quantitative and qualitative terms. (MH)

Final Report: A Study of the Communicative Abilities of Disadvantaged Children

Duration of Research: February 1, 1967 - January 30, 1968

Supporting Agencies:

1) Office of Research and Evaluation, Project Head Start, Office of Economic Opportunity (OEO Contract No. 2402)

2) The Johns Hopkins University School of Medicine

Principal Investigator: Harry Osser, Ph. D.

Research Staff: Evangelos Afendras, M.A.
Sheldon M. Frank, M.D.
Robert Jordan
Marilyn Wang, M.A.
Farida Zaid, B.A.

Acknowledgements

We would like to acknowledge the support and encouragement given this research by Leon Eisenberg, M.D., Harvard University Medical School (formerly of the Johns Hopkins University School of Medicine), Ellin Kofsky, Ph.D., and Sonia Osler, Ph.D. (both members of the Cognitive Studies research unit of the Johns Hopkins University School of Medicine), and finally to the teachers of the various schools in the Baltimore area where we worked.

Introduction

The first encounter between a group of lower class children just entering school and a middle class teacher beginning her first day of teaching is one that engenders shock and surprise for both of them. Each is not ready for the other. We know little about the feelings of the children but somewhat more about those of the neophyte teacher, as she at least can be questioned. A common response is apparently one of dismay. The classroom behavior of lower class children is noisy and non-compliant. The teacher quickly and rapidly concludes that she is grossly unprepared to teach them.

A frequent theme that pervades the discussion of the teacher's problems in teaching lower social class children involves communication difficulties. Some teachers when questioned have suggested that to develop adequate comprehension of lower class Negro children's speech takes several months after the first contacts are made with these children. It is most likely that the children themselves have a serious problem in understanding the teacher. What are the sources of these mutual problems in communication? Bernstein (1961) has proposed that the middle class teacher has an elaborated linguistic code; whereas the lower class child has a qualitatively different linguistic code which Bernstein has called a restricted code. It may be that these contrasting linguistic styles are a significant factor in the communication problems that exist between lower class children and middle class teachers: it is also quite likely that

dialect differences are important.

There is very little information available on the particular characteristics of lower class children's speech that differentiate it from the speech of middle class teachers. The research in this report was designed so that the code and dialect features of young lower class Negro children's speech and that of middle class white children could be precisely described.

Study #1

The Young Child's Ability to Imitate and Comprehend Speech: A Comparative Study of Two Sub-cultural Groups

Introduction

There exists a large number of studies that have focussed upon the problem of sub-cultural differences in speech development; both Cazden (1966) and Raph (1966) have reviewed this literature. Most of these studies have concentrated on the analysis of the expressive speech differences of the groups compared. However, language ability involves much more than the ability to produce speech; it includes, for example, the important component of speech comprehension, which McNeill (1966) has suggested is a basic linguistic ability, which, when thoroughly investigated will give considerable insight into the processes of language acquisition. Chomsky (1964) maintains that an adequate evaluation of the child's general linguistic abilities must include the experimental sampling of his various linguistic capabilities, and he makes the point that unless this is done the child's competence may be significantly underestimated.

There have been some recent studies that have examined linguistic abilities other than free speech in comparing the linguistic performance of sub-cultural groups (Peisach, 1965; and Salzinger, 1967). The following study was designed to evaluate some aspects of the syntactic development of young children in two distinct sub-cultural subgroups, Negro Lower Class (NLC) and White Middle Class (WMC). Two linguistic tasks, namely an Imitation task and a Comprehension task, were specially developed for this purpose.

Method

Subjects. Sixteen lower class Negro children and 16 middle class white children were used in the experiment. There was an equal number of boys and girls in each group. The NLC children ranged in age from 4 years, 7 months, to 5 years, 3 months. All were enrolled in an eight-week summer Head Start program in Baltimore. The Ss were tested during the last two weeks of the program. Subjects were chosen from the 75 children enrolled in the program on the basis of age and sex and all had normal hearing.

The performance of these Ss on the Imitation and Comprehension tasks was compared to that of 16 WMC Ss selected on the basis of age and sex, from a larger sample of children tested previously (Osser, Frank and Wang, 1967). The WMC Ss ranged in age from 4 years, 9 months, to 5 years, 2 months.

Materials Used to Test Imitation and Comprehension. The Imitation and Comprehension tasks were designed to test the Ss' control over thirteen syntactic structures. According to Roberts' (1964) transformational analysis of English syntax, these structures all occur as a result of transformations performed on kernel sentences in the language. For each of the thirteen structures two sentences were constructed which were syntactically identical but semantically different. One set of sentences was designated as Form A, and the syntactically equivalent set, Form B. All of the resulting sentences were seven words in length and all used simple vocabulary. In addition, five sentences were used for practice. The test and practice sentences are presented in Table 1.

Table 1

Practice Sentences

1. He throws the ball against the wall.
   She throws the ball against the wall.

2. The boy is going to slide down.
   The girl is going to climb up.

3. The clown has a big hat.
   The clown holds the ball in his hand.

4. The cat has a little tail.
   The dog has big ears.

5. The bunny is eating a carrot.
   The mouse is eating the cheese.

Test Sentences

1. The boy is pulled by the girl.
   The girl is pulled by the boy.
   (Passive)

2. The boy is not on the chair.
   The girl is not on the chair.
   (Negative)

3. Happily the boy is kicking the ball.
   Sadly the girl is kicking the ball.
   (Inversion)

4. The boy throws the ball far away.
   The boy throws the ball up high.
   (Separation)

5. The boy is pulling the girl's hair.
   The girl is pulling the boy's hair.
   (Possessive)

6. The boy dries himself with a towel.
   The girl dries herself with a towel.
   (Reflective)

7. The boy climbs and another boy climbs.
   The boy slides and another boy slides.
   (Conjunction)

8. The little boy is flying the kite.
   The big boy is flying the kite.
   (Adjective)

9. The boy who sits is very fat.
   The girl who sits is very fat.
   (Relative)

10. Mother does some sweeping with a broom.
    Father does some painting with a brush.
    (Nominalization)

11. The girl sees that the boy sits.
    The boy sees that the girl sits.
    (Transformation-
    Subordinate)

12. The boy does not wear a hat.
    The girl does not wear a hat.
    (Do and Negative)

13. The boy kicks the ball to her.
    The girl kicks the ball to him.
    (Transformation-
    Object)

When the test sentences had been completed, a corresponding set of pictures was designed. For each of the thirteen structures, a set of three pictures was made. Two of the pictures were representations of the two test sentences, and the third was a neutral picture. Within each set of three pictures, the same visual elements were present in each picture, but their spatial relationships differed. Therefore, the relationships between the elements in the picture provided the only cue for correct choice. In addition, the pictures were so designed that in order to respond correctly S must understand the meaning of the entire sentence. The Ss could not choose the correct picture on the basis of a partial decoding of the test sentence, except by chance.

The pictures were 4"x6" drawn in black ink on white paper, and were laminated with a heavy plastic. The figures in the pictures were designed to be as racially neutral as possible. One set of pictures is illustrated in Appendix 1.

Imitation Task. All Ss were tested individually in a private room. The instructions to S were as follows: "Listen very carefully. I am going to say something and when I am finished I want you to say just what I said." The five practice sentences were then presented to S, one at a time, to ensure that he understood the task. The 26 test sentences were then presented, one at a time, in a different random order for each S. If S failed to respond to the test sentence it was repeated once. The entire session was tape recorded.

Comprehension Task. On each trial three test pictures were placed before S, with the position of the correct picture varying randomly from trial to trial. The S was instructed by E as follows: "I am going to say something about one of the pictures and when I am finished I want you to

look at all of the pictures and find the one I am talking about." The S was given five practice trials with five different sets of pictures in order to ensure that he understood the task. The test sentences were then presented one at a time in a random order.

If S failed to respond to the test sentence within 10 seconds, the sentence was repeated. This was done to avoid errors due only to forgetting the test sentence. Occasionally Ss made more than one response on a given trial, or chose a picture without first looking at all of the pictures. In these cases the pictures were reshuffled, placed in front of the child again and the trial repeated. Only S's final choice was retained for analysis.

Both tasks were given in the same session and the order of the tasks was counterbalanced across Ss.

Methods of Analysis

Imitation Errors. Tape recordings for all Ss were transcribed independently by two judges, one Negro and one White. Points of disagreement were resolved by a third judge. All hesitation phenomena were then removed from the transcripts. These included vocal segregates (um, uh), repetitions (the, the boy), and self-corrections (the boy...the girl). When self-corrections occurred, the final version was retained for analysis.

Since the Imitation task was designed to test S's control over specific syntactic structures, it was the correct imitation of these structures which was of interest. For each of the thirteen sentence types certain words were designated as critical for correct imitation. The words so designated compose the Critical Structure. For example, in the sentences testing the child's acquisition of the relative clause, the words "who sits" comprise the Critical Structure.

All deviations from perfect repetition of the entire sentence were tabulated and classified according to the following scheme:

1. Omission of noun or verb inflection
2. Omission of word
3. Change in tense of verb
4. Change in number of noun or verb
5. Morphological error (hisself for himself)
6. Word substitution, same part of speech
7. Word substitution, different part of speech
8. Importation of word
9. Transposition of word order within sentence

Each S received two Imitation error scores. 1) Critical Structure Error score (CSE)--the number of sentences in which the S made at least one error (other than an error in category 6) in the Critical Structure. Errors in category 6 were not scored since they do not alter the syntactical structure of the sentence. This score operationally defines the number of sentences incorrectly imitated. 2) Total Error Score (TE)--the total number of deviations from perfect imitation of all the sentences. This score is a more comprehensive measure of the child's overall performance on the Imitation task.

Comprehension Errors. A Comprehension Error score (CE) was tabulated for each S and was based simply on the number of incorrect choices made in the Comprehension task.

Results

Differences Between the Two Groups. Mean CSE, TE and CE scores for the two groups are presented in Table 2. T-tests for these differences were respectively: 6.78, 7.95 and 4.11, df = 30. All of these differences are significant at the level p .002, two-tail test.

Table 2

Mean Errors on Imitation and Comprehension		
	WMC	NLC
CSE	2.06	11.00
TE	7.31	31.14
CE	1.67	6.00

Presumably, many of the errors made by the NLC Ss were attributable to the fact that the test sentences were encoded in standard English, i.e., a dialect that is structurally deviant from their own dialect (Stewart, 1965; Loban, 1966; Afendras and Osser, 1967; Labov and Cohen, 1967). Although the influence of dialect factors cannot be directly assessed for the Comprehension task, there is evidence for the role of dialect in the Imitation task. Responses of the NLC Ss clearly indicated that they were responding to the Imitation task with sentences which were encoded in dialect. Since error scores were based on a standard English model, all such dialect-bound errors inflated the error scores for the NLC Ss.

To make the comparison between the WMC and the NLC Ss as fair as possible, all relevant information about the non-standard dialect of the Negro children was taken into consideration. Loban (1966) has observed several categories of deviations that have a very high frequency for Negro children and a very low frequency for White children. These categories include: 1) Absence of the third person singular marker -s on the present tense of verbs; 2) Omission of the verb "to be"; 3) Omission of auxiliary

verbs; and 4) Non-standard use of verb forms. Examples of all these types of deviations did occur frequently in the imitations of the Negro Ss, but extremely rarely in those of the White Ss.

Taking these dialect variations into consideration, plus one other not mentioned by Loban, but frequent in our own data, viz., elision of the possessive marker -s, sharply reduces the CSE and TE scores of the NLC Ss and does not alter the scores for the White Ss. The adjusted means are shown in Table 3. T-tests for the differences are reduced to 2.79 and 3.77, df = 30, p .002, two-tail test. Thus, even with consideration given to dialect differences, the CSE and TE scores remain significantly higher for the NLC group.

Table 3

Mean Errors Adjusted for Dialect Differences		
	WMC	NLC
CSE	2.06	4.56
TE	7.31	17.56

The difference is further emphasized by comparing the distribution of all Imitation errors across the categories outlined above. (See Table 4.) As expected, the categories corresponding to known dialect variations show a large number of errors for the NLC group and virtually none for the WMC group. Yet in all of the remaining categories but one, the NLC Ss made more errors than the WMC Ss. Although some of these differences might reflect unidentified dialect variations, it is difficult to accept this as an explanation for all of the differences, particularly in those instances where the WMC group made many errors.

Table 4

Errors in Imitation		
Category	WMC	NLC
1. Omission of inflection		
*a. possessive		22
*b. third person singular -s	4	144
c. other		3
2. Omission of word		
a. article	25	99
b. noun		2
c. verb		7
d. auxiliary		
*i. be	3	51
ii. do		13
e. adverb	1	6
f. intensifier		3
g. relative pronoun	2	5
h. subordinator	6	4
i. preposition		1
3. Change in tense	4	13
4. Change in number	4	21
5. Morphological error	7	16
6. Word substitution, same part of speech	40	41
7. Word substitution, different part of speech	13	27
8. Importation of word	8	17
9. Transposition of word order		3
Total	117	498

* Categories reflecting known dialect variations

Differences Between the Imitation and Comprehension Tasks. Within each group a comparison was made between the mean CSE and CE scores in Table 2. T-tests for these differences for the WMC and NLC groups respectively were .94, df = 15, NS, and 4.91, df = 15, p .002, two-tail test. The significance of the difference for the NLC group is misleading, however. If the adjusted CSE score from Table 3 is used, t = -.11, df = 15, NS. Thus, the significant difference obtained using the standard scoring method may have been an artifact of that method. However, since it was not possible to adjust the CE scores for the dialect-bound errors, the comparison is at best suggestive.

In addition to looking at the mean level of performance of both groups on the two tasks, a more detailed analysis of the performance of the individual Ss was also carried out. If the Imitation and Comprehension tasks were, as the group means would suggest, of comparable difficulty within each group, and if they tapped the same or highly correlated linguistic abilities, two things should follow. First, a single S's performance on one task should be highly correlated with his performance on the other task. Second, sentences which proved difficult on one task should be of comparable difficulty on the other task.

The first of these hypotheses was tested by computing the Spearman Rank Order Correlation for performance on the two tasks, using the CSE and CE scores as indicators of performance. For the WMC Ss, rho = .26, NS; for the NLC Ss, rho = .55, p .05. Thus, there was some consistency in a single S's performance on both tasks in the NLC group but not in the WMC group.

The second hypothesis was tested by computing a difficulty score for each of the thirteen structures, based on the number of errors

(CSE or CE) made to that structure by all Ss, and computing the correlation between difficulty scores for the two tasks. For the WMC group, rho = .07, NS; for the NLC group, rho = -.07, NS. Thus, in neither group was the difficulty of a single structure preserved across tasks.

In spite of the lack of a difference in group mean performance on the two tasks, it cannot be concluded that the two tasks are equivalent. Rather, this result seems to be the outcome of a complicated set of interactions between subject task and structure in the WMC group and between task and structure in the NLC group.

Reliability of Error Scores. An estimate of the reliability of the error scores for the two tasks was computed by summing errors on Form A sentences and Form B sentences separately. Rank Order correlations for these scores are presented in Table 5.

Table 5

Reliability Coefficients Form A vs. Form B		
	WMC	NLC
CSE	.52**	.81*
TE	.78*	.87*
CE	.70*	.61*

* p .01 ** p .05

Discussion

By every indicator of performance that was used to examine the children's responses on the two tasks, the WMC group had significantly better scores than the NLC group. A crucial question is "What role did the NLC children's non-standard dialect have on their performance?" The

answer to this question is quite clear in the case of the Imitation task and much less clear in the case of the Comprehension task. Evidence of changes in the imitated sentences, by the NLC group, that are clearly contingent on the child's dialect can be observed throughout the data. For example, the sentence "The boy is pulling the girl's hair" is typically recoded as "(The) boy pulling the girl hair." Both the verb "be" and the possessive morpheme -s, (and often the initial article), are omitted. At least two of these omissions are predictable on the basis of our knowledge of the structural differences between the two varieties of English.

Thus, there is abundant data that the NLC Ss recoded the test sentence in the Imitation task. This result runs counter to that of Fraser, Bellugi and Brown (1963) who suggest that in imitating a sentence the child does not process it through his own meaning system. Their conclusion was based on the finding that children made fewer errors on the Imitation task and that they could parrot a sentence in imitation even though they could not identify the correct picture for that sentence on the comprehension task. Thus, imitation can occur without comprehension. However, the data for the NLC group suggest that whether comprehension occurs or not, the imitation response will tend to conform to the dialect which is familiar to the child even when the test sentence is not encoded in that dialect. This is evidence that the sentence is processed through some structural component of the child's linguistic system, if not his meaning system.

Even with dialect variations taken into account for the Imitation data, it was still the case that the NLC group made more errors on this task than the WMC group. Another possible explanation for this difference

might be that the NLC Ss performed more poorly because their memory span is shorter and the Imitation task requires the subject to temporarily store and then retrieve linguistic information. If this were the case, the difficulty of any given structure, as measured by CSE scores, should be correlated with the length of the Critical Structure, since longer Critical Structures put more of a load on memory and increase the probability that an error will occur. However, this was not the case. In neither group was the CSE difficulty score correlated with the length of the Critical Structure (WMC, rho = .11, NS; NLC, rho = .44, NS). Thus, in neither group can CSE scores be attributed solely to the length of the Critical Structure.

The differences between the groups on both the Imitation and Comprehension tasks strongly suggest that the NLC group's control over some common syntactic structures in standard English is markedly inferior to that of the WMC children. In addition, insofar as the corrections for dialect were accurate for the Imitation responses, it can be concluded that the NLC Ss show less control over functionally equivalent structures in their own dialect than the WMC Ss.

Moreover, the use of two tasks to evaluate Ss'control over specific syntactic structures was shown to provide more information than one task alone. Table 6 shows the number of children reaching a criterion of two correct responses on each structure for each task separately, and the number of children reaching a criterion of two correct responses on either the Imitation or Comprehension task. Predictably, more children from both groups demonstrated control over the structures when information from both tasks was considered, than would have appeared to be the case if only a single kind of information had been obtained. This fact

is directly related to the finding that the difficulty of a structure is dependent, to some extent, on the task used to evaluate it.

Table 6

Correct Responses by Structure
(16 Subjects per Group)

	Imitation*		Comprehension		I or C	
	NLC	WMC	NLC	WMC	NLC	WMC
Passive	13	14	3	10	14	15
Negative	16	16	13	16	16	16
Inversion	9	10	9	14	13	15
Separation	14	14	9	11	15	14
Possessive	4	16	14	16	14	16
Reflexive	2	8	15	15	15	16
Conjunction	13	16	4	15	15	16
Adjective	16	16	14	15	16	16
Relative	10	13	9	15	12	15
Nominalization	9	15	14	15	15	16
Subordinate	6	13	10	15	11	16
Do + Negative	12	13	12	16	15	16
Object	16	16	10	14	16	16

* For Negro children, known dialect differences taken into consideration.

These results have implications for the education of NLC children, since the language of instruction in the classroom in standard English, and often a teacher's expectations of their level of linguistic development in this dialect is unrealistic.

Study #2

A Comparative Study of the Syntax of Lower and Middle Class Children

A sample of free speech was collected from 20 lower class Negro 5-year-olds, and an equal number of white middle class children of the same age. The general mode of anlaysis that was used was tranformational grammar (Roberts, 1964). The purpose of the study was to discover any possible linguistic code variations existing between children of different sub-cultural backgrounds. In particular, an attempt was made to specify the syntactic structures that are common to both groups, as well as those structures that are unique to each group.

Some differences were observed between the two groups with respect to the range of syntactic structures they each exhibited. It can be observed from Table 1 (p. 20) that the white middle class group was superior in performance on 14 structures, and lower class Negro group on 5 structures.

Another type of analysis was used to uncover differences between the two groups in their use of specific types of complex sentences. The results are included in Table 2.

Table 2

Percentage of Children Using Various Types of Complex Sentences

	Structural Type	WMC	NLC	Example
1.	K*	6.1	9.6	The dog
2.	K* + op K unit	3.7	7.9	The black dog
3.	K* + T	8.6	8.5	The dog who has fleas
4.	K* + T + op K unit	4.4	4.3	The black dog who has fleas
5.	K	4.9	4.2	The dog bites
6.	K + op K Unit	7.2	12.7	The black dog bites
7.	K + T	17.4	13.4	The dog who has fleas bites
8.	K + T + op K unit	43.4	36.6	The black dog who has fleas bites
9.	Miscellaneous	4.3	2.8	

(K* Incomplete kernel sentence)

Categories 1-8 refer to sentence types of increasing complexity (see study #4 for a discussion of the development of this measure of complexity, and a fuller description of the structural types listed in Table 2). The white middle class children used proportionately more complex sentence types than the Negro lower class children; however none of the differences were found to be significant when Mann-Whitney U tests were applied.

Table 1

Percentage of children using each transformational type

Structure *	WMC		NLC	
	No. producing	Percentage	No. producing	Percentage
inversion	20	100	17	85
C-K	19	95	16	80
object	18	90	18	90
Conjunction phrase	18	90	16	80
do	17	85	18	90
negative	17	85	19	95
sub. II	17	85	10	50
sub. I	14	70	6	30
S-K	13	65	10	50
?	11	55	6	30
ind. question	11	55	2	10
relative	10	50	8	40
VT	9	45	9	45
passive	8	40	7	35
S-P	8	40	10	50
To-NP	8	40	4	20
possessive	7	35	5	25
yes/no	5	25	4	20
indirect object	5	25	5	25
interjection	4	20	5	25
there	3	15	3	15
apposition	3	15	9	45
separation	3	15	2	10

* These structures are defined in Roberts (1964)

Examples of Transformation Types Used by Both Groups

inversion	Mostly I play with cars.
C-K	The doggie is blowing bubbles in the bathtub and he's washing in the bathtub.
object	The cat has a bullet belt all around him.
Conjunction Phrase	I make an army of army men and an army of marines.
ao negative	This cat doesn't like the mouses.
sub_{II}	Once when my father strained his back he had to lie on the floor too.
sub_{I}	She doesn't think that everything's going right.
S-K	They just work, daddy goes to work.
?	What's this?
ind. question	That's why she's looking into there.
relative	That's a truck that's carrying toys.
VT	They have to send it away.
passive	It has to be trained in the circus.
S-P	She plays games, stuff like that.
To-NP	The girl's trying to open the door.
possessive	He's putting his stethoscope on the dog's heart.
yes/no	Is that a new one?
indirect object	I gave my friends a valentine.
interjection	Clunk! the dog comes off, too.
there	There is still more snow outside.
apposition	My sister Ann and my baby sister plays together.
separation	The father said let's go in.

Study #3

A Contrastive Study of the Dialects of White Middle Class and Negro Lower Class Children

An accurate description of linguistic code differences between children from different sub-cultural groups has to take dialect variations into account, otherwise the comparison is invalid. The non-standard Negro speaker will produce the sequence "The boy hair" instead of the standard "The boy's hair," so that if his speech were evaluated by referring to the rules of standard English, he would not be credited with the control of the possessive structure. This would be a completely erroneous interpretation, as the Negro does in fact provide evidence of his control of this structure in producing the sequence "The boy hair": the form is non-standard, but it is perfectly consistent with his grammatical system.

A detailed sketch of non-standard Negro dialect is obviously necessary before differences that have been observed between the linguistic performances of Negro lower class children and White middle class children can be accepted as real differences. The purpose of this study was to describe some of the features of non-standard Negro dialect. All speech of the Negro children in Study #2 was examined for the presence of non-standard syntactic variations. The speech of the White children in that study was analyzed for control purposes. These children only rarely exhibited non-standard phenomena, so that the following description relates essentially to the dialect features observed in the Negro children's speech.

1) Ommission of the copula (in various environments)

1. She combin' her hair.
2. Them playin'.
3. This a tiger an' that a bear.
4. They tryin' to hit the ball.

(However sentences with a mention of places did not usually omit the copula, e.g., "They're in a truck.")

2) Personal and Possessive Pronouns

1. Them playin' tennis.
2. An' the little girl, he put on a scarf.
3. Him climb up.
4. Him puts him foots in the water.

3) Genitive

1. That ball hit that dog head.
2. Didn't tell me go roun' Bill house.

4) Verb forms

1. While he been writin'.
2. He just go out the bathtub.
3. My brother Jim and my sister gone away.
4. When he be upstairs putting Junior to bed, I sneak and I put my clothes on.

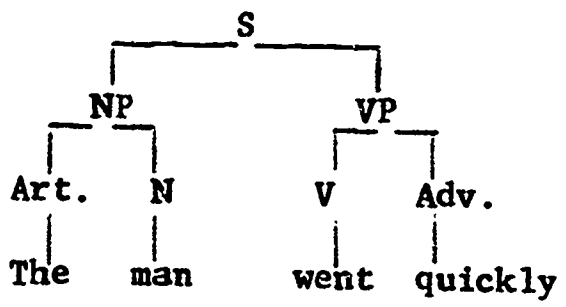
Study #4

A Psycholinguistic Model of Syntactic Complexity

Introduction

There have been several recent conceptualizations of the nature of syntactic complexity which have described it in quantitative terms. This has been done by referring to the depth of a sentence (Yngve, 1960), its node/terminal ratio (Miller and Chomsky, 1963), or its degree of nesting (Schlesinger, 1966). Each one of these approaches to syntactic complexity will be discussed briefly, and then a psycholinguistic model of syntactic complexity which is being developed will be presented.

Yngve (1960) has proposed a Push-Down-Store model of sentence generation, whose central feature is a temporary memory bank that can store only a finite number of symbols. Yngve's quantitative index of syntactic complexity, i.e., depth, refers to the amount of temporary memory needed to produce a sentence. For example, consider the production of the sentence: "The man went quickly." The tree diagram would be:



The speaker generates a portion of the sentence, and the rest of it must be kept in the memory store, so that when:

S is reduced to NP + VP, no other elements are kept in the store.

NP " " " Art. + N, 1 " element is " " " " (VP)

Art. " " " 'The', 2 " elements are " " " " (N, VP)

N " " " 'man', 2 " " " " " " " (Art., VP), etc.

For each of these transactions, the number of elements required to be put in the PDS is equated to d or depth at each point in the generation. The maximum value of d (in this case, 2) is considered to be D, or depth of the sentence. Any utterance can be given a value of from 0 to 7 on this scale of sentence depth. The maximum value of 7 is consonant with the results of psychological studies of the immediate memory span (Miller, 1956).

Schlesinger's (1966) notion of degree of nesting refers to any sentence of the form a b c ... n ... c' b' a', where a' completes a, b' completes b, etc., and all parts except a and a' are nested parts. The number of such pairs in addition to a and a' is equivalent to degree of nesting. Thus the simple sentence used above would have a degree of nesting (d.n.) of 0. The sentence, "I know the man who did it personally," has a d.n. of 1.

A third possible parameter of syntactic complexity, the node/terminal node ratio, is described by Miller and Chomsky (1963). To derive this measurement, the total number of nodes in the tree diagram must be counted and divided by the number of those markers below which no further branching occurs. For example, in the tree diagram above, the nodes to be considered are S, NP, VP, Art., N, V, Adv. The last four of these are terminal nodes. Thus the node/terminal node ratio here is 7/4 or 1.75.

Purpose and Approach

The goal has been to develop a model of syntactic complexity that could be used in psycholinguistic research. In particular, the aim has been to provide both a quantitative and qualitative description of language performance, so that comparisons can be made between individuals at different stages of language development. In addition, it is expected that such a system might also differentiate individual styles of speech. As will be seen in the following two sections, the theoretical basis of this model is

a modified transformational grammar approach (Roberts, 1964; and Chomsky, 1965), to which a quantified operational analysis has been applied.

Transformational Theory

From the point of view of transformational generative grammar (Chomsky, 1965), every sentence can be described as either a kernel sentence, or as a more "complex" sentence which includes one or more transformations. This second category of sentences is viewed as being derived from kernel sentences by syntactic rules known as transformations. For example, the sentence, "The elephant has been a thief," is a kernel sentence, which when processed through the negative transformational rule becomes "The elephant has not been a thief" (NP & aux. & be & substantive ⟶ NP & aux. & not & be & substantive).

The first category, or kernel sentence, is defined as K ⟶ NP & VP. An underlying nucleus for each K has been devised by further analysis of its main components, viz., NP & VP. These have been subdivided into essential subcomponents (belonging to the nucleus), and optional[1] subcomponents. Some samples of the analysis follow:

Noun phrase analysis

NP ⟶
- Type 1) proper noun + (adj. - prep. phrase)
- or
- Type 2) indefinite pronoun + (adj. prep. phrase)
- or
- Type 3) (pre.-art.) + /̄poss. pron./̄ + /̄demon./̄ + (number) + (adj.) +(N + adj. - prep. phrase)

---

1 Material which is in parentheses or brackets. That in parentheses is thought to add significantly to the complexity, and is scored accordingly. That in brackets is not thought to add complexity sufficient to be quantitatively measurable in this system. A more refined system may attempt such measurement in the future.

Verb Phrase Analysis

VP → tense + (M) + (have and part.) + (be & ing) +

1) be + (intens.) + /substantive/ or /adv. p./
or
2) verbal

} + (intens.) + (Adv.)

To illustrate these procedures, the sentence, "That man from the office whom I saw last week is very sick," will be analyzed. Removing the transformations, the kernel sentence would be "That man from the office is very sick." The underlying nucleus of the kernel, consisting of its essential subcomponents, is "man is sick," where the NP is of type 3 and the VP of type 1. All other components of the sentence, outside of the nucleus, are optional subcomponents; these may (as in the example) or may not include "transformations."

Some recent work in linguistics (viz., Koutsoudas, 1966) has stressed the subdivision of a grammar into phrase structure rules (P-rules) and transformational rules (T-rules), the latter in turn being composed of obligatory and optional transformations. Among the benefits envisaged for this type of analysis is a more rigorous definition of the kernel sentence, which is defined as the result of application of P-rules and obligatory T-rules. No practical and comprehensive derivates of this approach have as yet been worked out. Should it prove useful, it is very likely that the operational analysis and other main tenets of the model presented here would remain applicable. For instance, whether certain transformations may be reconsidered to be P-rules, obligatory T-rules, or optional T-rules, the syntactic complexity should remain the same, the molecular basis being unchanged.

In experimental work on the analysis of children's free speech (Osser, Frank, and Wang, 1967), a large number of transformations have been encountered, some of which are described in Table 1. Roberts (1964) and Menyuk (1964) have each provided models for many of these transformations.

Table 1

Selected Types of Transformations Found In Children's Speech

Single-based transformations (derived from a single kernel)

	Transformation	Definition	Illustration
a)	T - inversion	NP & VP → {(Adverb) or (Conjunction) or (Interjection)}	The elephant has been a thief → Often the elephant has been a thief.
b)	T - yes/no	NP & aux. & be & substantive → Aux. & NP & be & substantive →	The elephant has been a thief → Has the elephant been a thief.[2]
c)	T - ?	NP & VP → NP & VP & question inflection	Has the elephant been a thief → Has the elephant been a thief?
d)	T - wh - adv.time	NP & VP adv. time → When & NP & VP	The elephant has been a thief today → When the elephant has been a thief.[3]
e)	T - negative	NP & aux. & x → NP & aux. and not & x	The elephant has been a thief → The elephant has not been a thief.

Double-based transformations (derived from two kernels)

f)	T - relative	Matrix: X & NP & Y	Yesterday we noticed the birds.
		Insert: Z & NP & W	On Friday the birds escaped from the zoo.
		Result: X & NP & who or which or that + Z & W & Y	Yesterday we noticed the birds, who on Friday escaped from the zoo.

---

2 The result of processing the K structure through (b) does not generate a question: the K structure has to be processed through both rules (b) and (c) to achieve this. Notice, however, that in colloquial speech (c) can be employed without (b), e.g., "The elephant has been a thief?"

3 To generate the standard question form "When has the elephant been a thief?", the sentence would have to be processed through rules (b) and (c).

Operational Analysis

It is important to describe and quantify the different logical operations involved in the process of encoding from the nucleus to the kernel and finally to the transformed sentence. The primary operation appears to be the creation of the subcomponents of the nucleus. Beyond this point, four kinds of operations seem to apply:

1) addition of a word or a phrase

2) subtraction of a word or a phrase

3) transposition of a word or a phrase

4) intonational change

These operations can be illustrated with reference to the sentence types in Table 1. In each of the single-based transformation sentences, the primary operation was the creation of an NP & a VP, here elephant and be thief respectively. The secondary operations can be exemplified as follows:

1) addition: not in (e)

2) subtraction: adv-time in (d)

3) transposition: aux. in (b)

4) intonational change: interrogative intonation in (c)

It has been assumed that these four types of operations represent logical steps of roughly equal difficulty. With the first three types of operation, the formation of the terminal string has been considered to be analogous to a row of building blocks, and the effort required to "lift" a component into the string, out of the string, or from one point of the string to another, to be of the same magnitude. Similarly it has been supposed that the significant (i.e., morphemic change in the

intonation of a sentence, and finally the primary operation of creating the NP & the VP are each of the same order of complexity as the other operations.

Types of analysis possible

The operational analysis described above permits both qualitative and quantitative treatments of a speaker's linguistic performance:

Qualitative analysis: some examples

1) Categorization of each terminal string: nucleus only, complete K, transformed sentence, or sentence with multiple transformations.
2) Categorization of kernels: simple, complex, or deviant form.
3) Description of specific transformations.
4) Classification of transformations: single-based or double-based.

Quantitative analysis: 1) Simple and complex kernels

In the quantitative analysis each sentence is allotted a number of units of complexity, the number being determined by the total of the logical operations involved in the generation of the sentence. One unit is given for each of the components of the nucleus (i.e., NP & VP) present, thus "That man is sick" (a simple kernel) gets a score of two. In the expanded sentence "That man from the office is very sick" (a complex kernel), an extra unit is given for each of the two additions "from the office" (adj.-prep. phrase) and "very" (intens.), so that this sentence has a total complexity score of four units.

Quantitative analysis: 2) Transformed sentences

Sentences which include transformations can be broken down operationally into combinations of the operations of addition, subtraction, transposition, and inflectional change. Each of these four types of change has been assigned one unit of complexity. Thus, the sentence "Sometimes I stand up" is given a score of 3 units of complexity (1 for NP, 1 for VP, and 1 for a transposition, i.e., the inversion transformation). To streamline syntactic analysis of material, the list of transformations that have developed includes the usual score of that transformation, based on this operational analysis, e.g., T-neg. can be assumed to have 1 unit, and T-subord., 4 units, providing no unusual deletions or additions are made.

Comparison of 4 different approaches to syntactic complexity

Below are 3 sentences which have been selected at random from each of 2 speakers in one of the experimental groups of pre-school children. The speech material was elicited by putting simple questions to the child. To these sentences 4 schemes of syntactic complexity analysis have been applied in turn: depth, degree of nesting, node/terminal node ratio, and the new system which has been presented above.

The sentences to be analyzed are the following:

Speaker H.E.: 1) We sing songs and play.

2) I never played only one time.

3) They sneak in.

Speaker S.S.: 1) We sit on the circle and stand on the circle.

2) You set them up on the floor because they can't stand on the rug.

3) Did you see Tinkerbell die?

By the method of depth (see above, p. 2), the following scores are obtained:

H.E.	1) 3	2) 3	3) 4	Average: 3.3
S.S.	1) 4	2) 4	3) 3	Average: 3.7

The degree of nesting analysis (see p. 2) yields these results:

H.E.	0 for all three sentences
S.S.	0 for all three sentences

Node/terminal node ratios for the same sentences are:

H.E.	1) 1.8	2) 1.6	3) 2.0	Average: 1.8
S.S.	1) 1.7	2) 1.8	3) 1.8	Average: 1.8

(See Appendix 1 for tree diagrams by which these figures are derived.)

Finally by the present approach of tabulating units found in kernel and transformations, the following scores are computed:

H.E.	1) 4	2) 5	3) 2	Average: 3.7
S.S.	1) 6	2) 12	3) 10	Average: 9.3

(See Appendix 2 for analysis.)

Discussion

In the group of subjects under investigation (Osser, Frank and Wang, 1967), as well as in the sample sentences cited above, the present system corroborated common-sense notions of relative complexities of syntax, whereas the three other methods generally did not.

In addition, it has been observed that both the average sentence complexity scores, and the number of different T's used, as well as the proportions of double-based T's generally increases with age, which also

gives support to the validity of the system. A few of the persistent problems will be discussed below.

I. The problems of depth and nesting

There are no direct provisions for depth or nesting in this system. However, to some extent, the depth or degree of nesting of a sentence will be accounted for indirectly by the large number of operations required to produce sentences which rank high on these two measures. Also, it is felt that in part a question of style and usage has a role in explaining the difficulty of certain of such sentences, at least for a degree of nesting of two or less. Consider the following sentences:

a) John whom June whom Paul prefers detests loves Mary.

b) Paul prefers June and she detests John and he loves Mary.

c) Mary is loved by John, who is detested by June, who is preferred by Paul.

The style/usage question might well play a role in comparing (c) with (a). Rather than the relative difficulty of (a) being explained by inherently greater complexity of nested relative over passive relative, it may be that in most people's common style there is seldom an occasion for the very nested or deep sentence, and that were it of common usage, its exigencies would appear less. In other words, the difficulties of "holding" a part of a sentence or message in memory may be quite resolvable by practice; it would seem that native German speakers do this with relative ease. This problem is obviously not completely solved, as can be observed in the following. Common sense would rate (a) as most complex, followed by (c) and then (b). Degrees of nesting are 2 for (a) and 0 for (b) and (c). By depth analysis, (a) and (b)

have depth scores of 4, (c) has depth of 3. Under this scheme, (b) receives 8 units, (a) 12, and (c) 25. A possible method to resolve this implausible result (with respect to (a) vs. (c)) would be a modification of the operation addition. It may well be that adding an element (e.g., "whom Paul prefers" in (a), "and she detests John" in (b), etc.) in the middle of a kernel component is a more complex encoding task than adding an element at either end. To return to the somewhat crude building block analogy, it does seem easier to add the blocks to one end of the row, than to shove aside some blocks in mid-row for addition. This will have to be elaborated, and it must be mentioned that the work of at least one investigator (Coleman, as cited in Schlesinger (1966, p. 123)) supports the increased complexity of deep sentences. However, Schlesinger (1966) in his own experimental work found no significant difference in comprehension of nested and unnested sentences.

II. Usage

As alluded to above, there are situations where the T seems either more or less difficult than the number of operations would predict, and that the usage factor appears to play a role. For example, the sentence "I am going to do this" seems a priori much more difficult than its common usage by small children would indicate. Most likely, as in word (and contraction) etymologies, the component steps are somehow obliterated with common usage in such T's, and they are learned more as a unit than in a series of steps. Indeed in the case of "going to," "want to," "have to," one is tempted to treat them as single syntactic entities, i.e., as Auxiliaries.

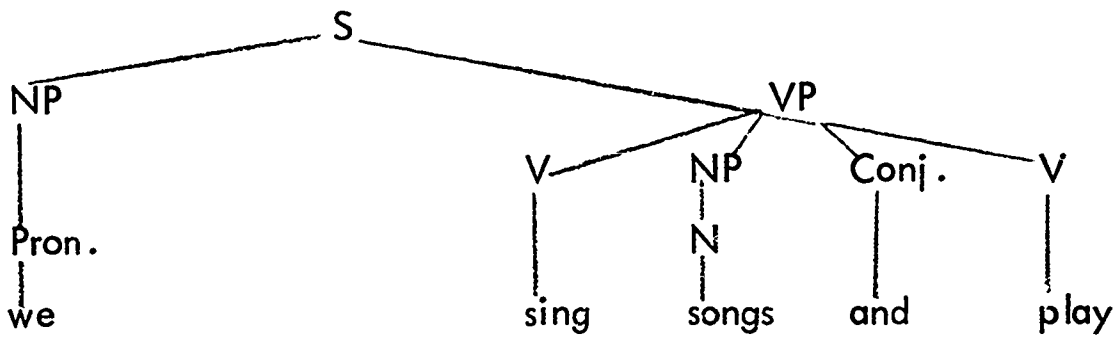
III. Total number of T's

There is some problem as to how many T's to list. Probably a certain amount of arbitrariness in inevitable. What is a T and what is an individual or a deviant pattern can be a matter both of style and of dialect--in both cases a law will have to be followed. When the dialect has a genuinely new T vs. a modification of the "standard" T, this is a further problem--it need not be of concern here.

Appendix 1

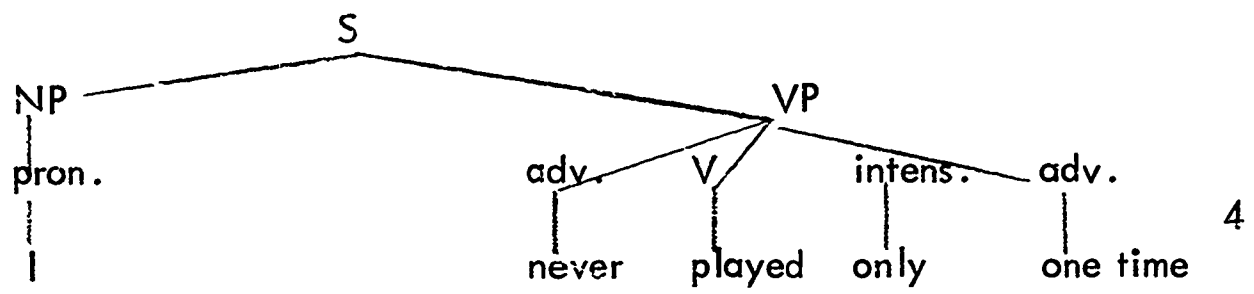
Analysis of Sample Sentences by Node/Terminal Node Ratio

H.E. 1)



Nodes: 14, terminal nodes: 5 ratio = 9/5 = 1.8

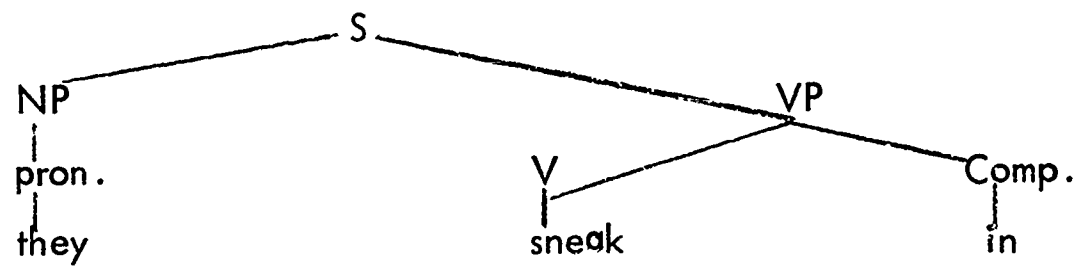
H.E. 2)



4

Nodes: 13, terminal nodes: 5 ratio = 8/5 = 1.6

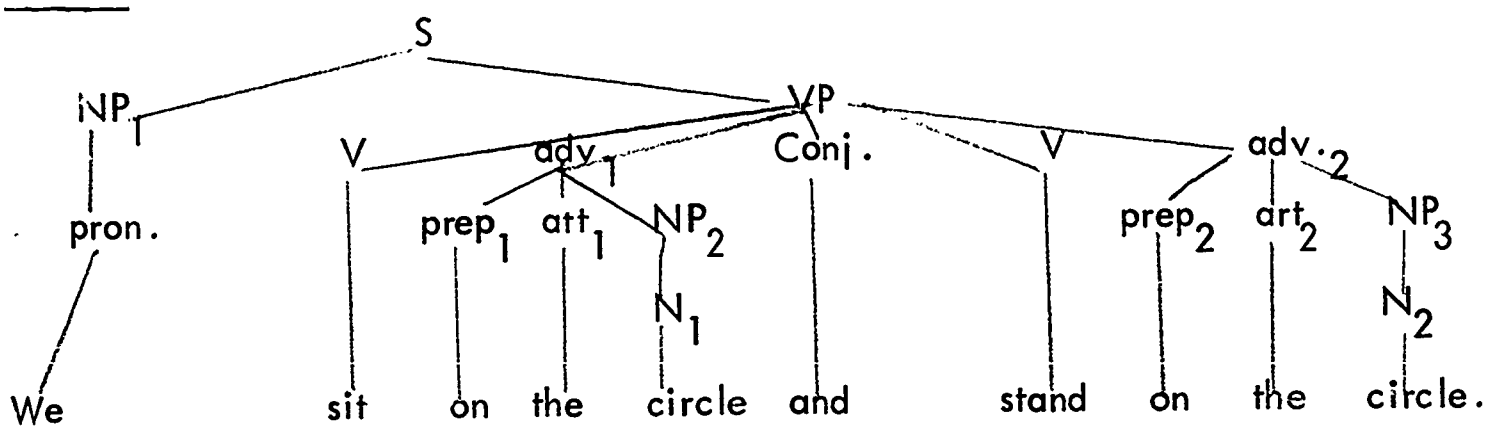
H.E. 3)



---

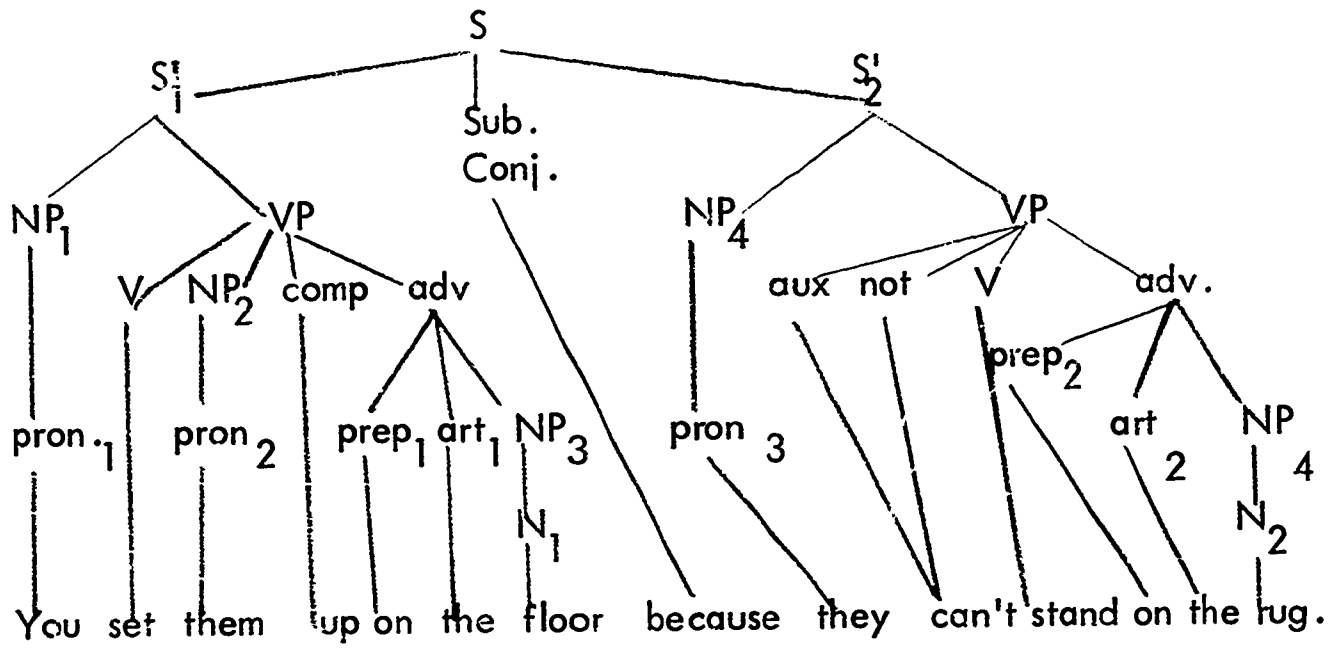
4 This has been treated as the equivalent of the simple adverb "once."
For non-complex NP's and prepositional phrases used as adv, it is believed that this is justified.

S.S. 1)



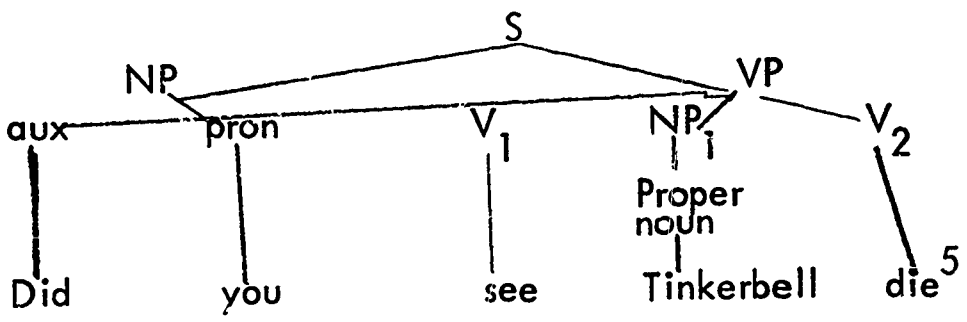
Nodes: 17 terminal nodes: 10 ratio = 17/10 = 1.7

S.S. 2)



Nodes: 27 terminal nodes: 15 ratio = 27/15 = 1.8

S.S. 3)



Nodes: 14 terminal nodes: 5 ratio = 9/5 = 1.8

---

5 In this and the previous sentence, it may be that the method of deriving the ratio should depend on 2 trees -- i.e., for each kernel of the double-base sentence. This might explain the failure of tree diagrams to illustrate the deep (vs. surface) aspect of complex sentences. However, no such method appears in the literature, and use of the 2 trees has not altered the scores significantly.

Appendix 2

Analysis of Sample Sentences by the Newly Developed Syntactic Complexity System

H.E. 1)	Kernel	
	nucleus (we sing songs)	2 units
	optional	0
	Transformations	
	T - conj - P	2
	Total	4 units
2)	Kernel	
	nucleus (I played)	2
	optional - adv - M(never)	1
	adv - T (one time)	1
	intens (only)	1
	Transformations	0
	Total	5 units
3)	Kernel	
	nucleus (they sneak in)	2
	optional	0
	Transformations	0
	Total	2 units
S.S. 1)	Kernel	
	nucleus (we sit)	2 units
	optional - adv - P (on the circle)[6]	1
	Transformations	
	T - conj - P	2
	Optional element added to transform adv - p - (on the circle)	1
	Total	6 units

6 Prepositional phrases thus far have been treated as a single unit, equivalent to one-word adverbs: obviously more complex examples, such as "on the 21st glorious star-studded circle" would have to account for optional items. (See Footnote 1)

S.S. 2) Kernels

S_1	(matrix) - nucleus - (you set up them)	2
	optional - adv - P - (on the floor)	1
S_2	(insert) nucleus (They stand)	0[7]
	optional /aux (can)	1
	/adv - P (on the rug)	1

Transformations[8]

T - sub_2	4
T - VT	1
T - reg.	1
T - obj.	1
TOTAL	12

3) Kernel

nucleus (you saw Tinkerbell)	2
optional	0

Transformations

T - VT_{to}	4
T - do	1
T - yes/no	1
T - ?	1
T - VT	1
TOTAL	10

---

7 Included in transformation (T - Sub_2)

8 Order in which transformation has been performed has not been thought to be crucial.

References

Afendras, E., and Osser, H. A contrastive study of the dialects of White middle class and Negro lower class children, 1967 (Mss.).

Bernstein, B. Social class and linguistic development: A theory of social learning. In A. H. Halsey, J. Floud, and C. A. Anderson (Eds.), Education, economy and society. Glencoe: The Free Press, 1961, pp. 288-314.

Cazden, Courtney B. Some views on subcultural differences in child language. Merrill-Palmer Quarterly, 1966, 12, 185-220.

Chomsky, N. Formal discussion of "The development of grammar in child language" by W. Miller and Susan Ervin, Society for Research in Child Development Monographs, 1964, 29, 35-39.

Chomsky, N. Aspects of the theory of syntax. Cambridge: M. I. T. Press, 1965.

Fraser, C., Bellugi, Ursula, and Brown, R. Control of grammar in imitation, comprehension and production. Journal of verbal Learning and verbal Behavior, 1963, 2, 121-135.

Koutsoudas, A. Writing transformational grammars, New York: McGraw-Hill, 1966.

Labov, W., and Cohen, P. Systematic relations of standard and non-standard rules in the grammar of Negro Speakers, 7th Project Literacy Conference. Cambridge, 1967.

Loban, W. Problems in oral English. Champaign: National Council of Teachers of English, Research Report, 1966, No. 5.

McNeill, D. Developmental psycholinguistics. In F. Smith and G. A. Miller (Eds.), The genesis of language. A psycholinguistic approach. Cambridge: M. I. T. Press, 1966, pp. 15-84.

Menyuk, Paula. Syntactic rules used by children from preschool through first grade. Child Development, 1964, 35, 533-546.

Miller, G. A. The magical number seven, plus or minus two: Some limits on our capacity for processing information. Psychological Review, 1956, 63, 81-97.

Miller, G. A., and Chomsky, N. Finitary models of language users. In R. D. Luce, R. Bush, and E. Galanter (Eds.), Handbook of Mathematical Psychology, Vol. 2, Ch. 13. New York: Wiley, 1963.

Osser, H., Frank, S. M., and Wang, Marilyn. The communicative abilities of preschool children, 1967 (Mss.).

Peisach, E. Children's comprehension of teacher and peer speech. Child Development, 1965, 36, 467-486.

Raph, Jane B. Language development in socially disadvantaged children. Review of Educational Research, 1965, 35, 389-400.

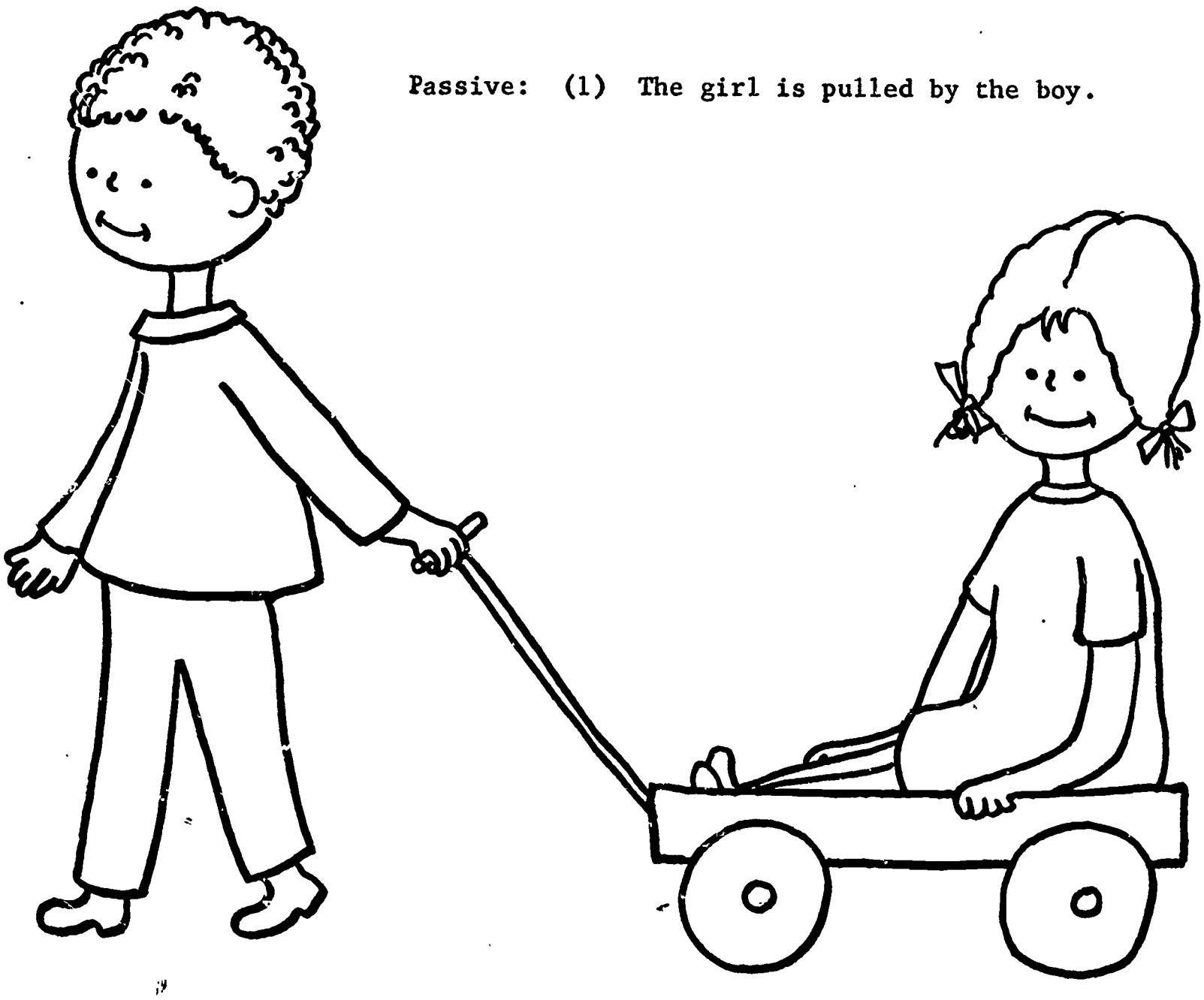
Salzinger, S. Some implications of methodology for the development of the field of child language, Paper presented at S.R.C.D. meetings, N.Y.C., 1967.

Schlesinger, I. M. Sentence structure and the reading process. The Hague: Mouton, 1966.

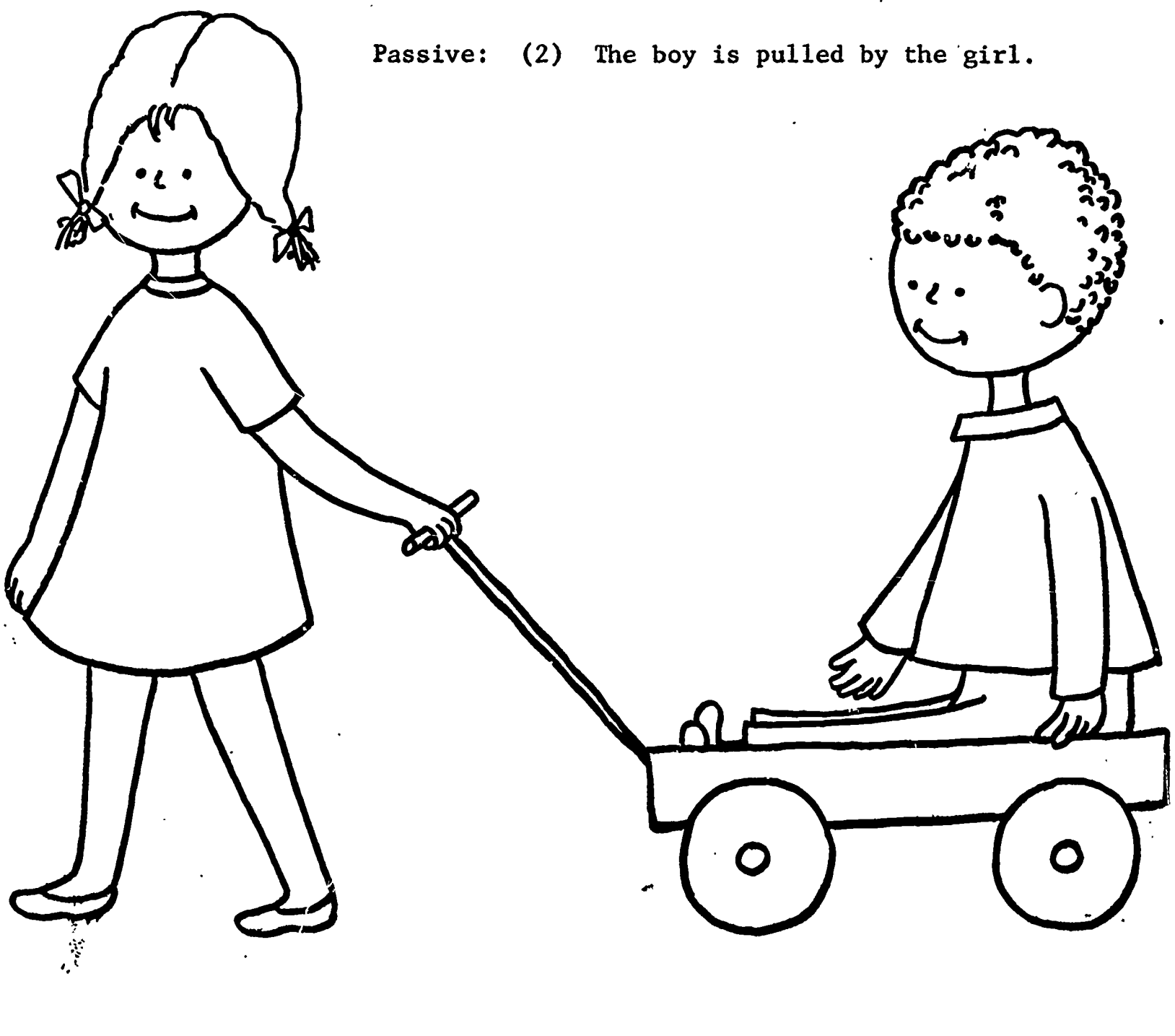
Stewart, W. Non-standard speech and the teaching of English. Washington, D.C.: Center for Applied Linguistics.

Yngve, V. A model and a hypothesis for language structure. Proceedings of the American Philosophical Society, 1960, 104, 444-466.

Passive: (1) The girl is pulled by the boy.



Passive: (2) The boy is pulled by the girl.



Passive: (3) Neutral